DOCUMENT RESUME

ED 218 099 SE 037 899

AUTHOR Kleinbaum, David G.; Kleinbaum, Anna
TITLE Adjusted Rates: The Direct Rate. Applications of Statistics to Demography. Modules and Monographs in Undergraduate Mathematics and Its Applications Project. UMAP Unit 330.
INSTITUTION Education Development Center, Inc., Newton, Mass.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 79
GRANT SED-76-19615-A02
NOTE 43p.

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Answer Keys; *College Mathematics; Demography; Higher Education; Instructional Materials; *Learning Modules; *Mathematical Applications; *Problem Solving; *Statistics; Supplementary Reading Materials

ABSTRACT

The document notes that in health and medical research the need to compare rates or proportions for some event can often occur. It is the module's goal to illustrate the conditions necessary for recognizing potentially misleading situations, and to show how to cope statistically with such situations. The four sections show: 1) when to adjust rates; 2) what adjusted rates do and how to compute the directly adjusted rate; 3) how to interpret adjusted rates; and 4) how to compute and interpret directly adjusted rates when adjusting for factors other than age. Both exercises and a posttest are presented, with answers to both provided. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED218099

umap

UNIT 330

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

ADJUSTED RATES: THE DIRECT RATE

by David G. Kleinbaum and Anna Kleinbaum

APPLICATIONS OF STATISTICS TO DEMOGRAPHY

edc/umap /55chapel st./newton,mass.02160

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

National Science Foundation

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

by

David G. Kleinbaum
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514

and

Anna Kleinbaum
307 McCauley Street
Chapel Hill, NC 27514

TABLE OF CONTENTS

INTRODUCTION 1

1. WHEN TO ADJUST RATES 2
2. WHAT ADJUSTED RATES DO AND HOW TO COMPUTE THE DIRECT RATE . . 9
3. INTERPRETING ADJUSTED RATES 20
4. REVIEW USING A CONFOUNDING FACTOR OTHER THAN AGE 21
5. POST TEST 23
6. RECOMMENDED FOLLOW-UP ACTIVITIES 27
7. APPENDIX: A STANDARD MILLION 27
8. ANSWERS TO REVIEW SECTION QUESTIONS 28
9. ANSWERS TO IN-TEXT QUESTIONS 1-20 30
10. ANSWERS TO POST TEST 32
11. REFERENCES 35

Intermodular Description Sheet: UMAP Unit 330

Title: ADJUSTED RATES: THE DIRECT RATE

Authors: David G. Kleinbaum
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514

and Anna Kleinbaum
307 McCauley Street
Chapel Hill, NC 27514

Review Stage/Date: III 9/28/79

Classification: APPL STAT/DEMOGRAPHY

Prerequisite Skills:

You should be able to work through this package successfully without having formally taken a first course in biostatistics, although you will likely be more familiar with the terminology used and have more insight into the subject matter of this package having had such a course. Regardless of your formal background in biostatistics, you should have the following knowledge as prerequisites:

1. Experience in reading and constructing tables and graphs.
2. Ability to make accurate mathematical calculations either by hand or by calculator.
3. Ability to identify and compute a rate or proportion (p).
4. Ability to convert a rate in any base to a proportion and vice versa.
5. Ability to define the following terms (refer to these definitions if you have trouble later):

   a. *population* - the totality of people defining a group of interest at the time of interest.
   b. *test population* - a population for which you have a question of interest (for which you wish to test something).
   c. *overall rate* - any rate describing or summarizing experience in an entire population (as opposed to a specific subgroup) for some characteristic of interest.
   d. *crude rate* - an overall rate defined by the formula: number in entire population with characteristic of interest during the time of interest/total number in entire population during the time of interest.
   e. *specific rate* - a rate for a specific subgroup of a population of interest (example: age group 10-20 years).
   f. *distribution* - a table or graph which shows the (relative) frequency of persons in a population distributed into non-overlapping categories of a variable of interest.
   g. *population-at-risk* (PAR) - the number of people in a population used in the denominator of a rate.
   h. *a standard million* - a population (e.g., U.S. 1960) whose numbers in specific categories have been changed so as to total one million while remaining in the same relative proportions.*

---

*If you would like more elaboration with an example, turn to Section 7, Appendix.

Output Skills:

This programmed instructional package will show you when and why you need to adjust rates and how adjustment by the direct method is done. You will learn how to adjust for factors other than age and how to interpret the results. Health-related examples are given and you are encouraged to learn by solving problems. When you have completed this program you will be able to:

1. State the conditions necessary for rate adjustment.
2. Use these conditions to evaluate whether such adjustment is appropriate.
3. Given the necessary basic information, compute the direct adjusted rates for two populations of interest.
4. Interpret the results of your computations as to the comparison of the overall mortality or morbidity experiences between the two populations of interest.
5. Compute and interpret adjusted rates when the confounding factor is a variable other than age.

Estimated Working Time: 1½ - 2½ hours.

Intended Audience: Health scienc_ students or professionals studying epidemiology and/or introductory biostatistics; students of demography.

Other Related Units:

© 1979 EDC/Project UMAP.
All rights reserved.

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT (UMAP)

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications which may be used to supplement existing courses and from which complete courses may eventually be built.

The Project is guided by a National Steering Committee of mathematicians, scientists, and educators. UMAP is funded by a grant from the National Science Foundation to Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research in the U.S. and abroad.

PROJECT STAFF

Ross L. Finney	Director
Solomon Garfunkel	Associate Director/Consortium Coordinator
Felicia DeMay	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Paula M. Santillo	Administrative Assistant
Donna DiDuca	Secretary
Zachary Zevitas	Staff Assistant

NATIONAL STEERING COMMITTEE

W.T. Martin	M.I.T. (Chair)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
Ernest J. Henley	University of Houston
William Hogan	Harvard University
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
R. Duncan Luce	Harvard University
George Miller	Nassau Community College
Frederick Mosteller	Harvard University
Walter E. Sears	University of Michigan Press
George Springer	Indiana University
Arnold A. Strassenburg	SUNY at Stony Brook
Alfred B. Willcox	Mathematical Association of America

UMAP STATISTICS PANEL

Roger Carlson	University of Missouri, Kansas City
Earl Faulkner	Brigham Young University
Tom Knapp	Rochester University
Peter Purdue	University of Kentucky
Judith Tanur	SUNY at Stony Brook
Richard Walker	Mansfield State College
Douglas A. Zahn, Chair	Florida State University

This module was developed under the auspices of the UMAP Statistics Panel. The Project would like to thank Douglas A. Zahn, Chairman, and, Judith Goldberg, Charles H. Goldsmith, Donald Guthrie, Duane A. Meeter, and Janet Wittes for their reviews of this unit.

This material was prepared with the support of National Science Foundation Grant No. SED76-19615 A02. Recommendations expressed are those of the author and do not necessarily reflect the views of the NSF, nor of the National Steering Committee.

ADJUSTED RATES: THE DIRECT RATE

INTRODUCTION

If you ever become involved in health or medical research, whether through actual experience or through reading the literature, you are very likely to encounter the need for a *comparison of rates or proportions* for some event or characteristic across different populations of interest. If these populations (e.g., different communities, treatment groups, exposure groups) are similarly constituted with respect to factors (such as age, sex, race) associated with the event under study, there would be no problem in comparing simple *crude* rates as they stand. However, if the populations are not similarly constituted, a straightforward comparison of crude rates may be misleading.

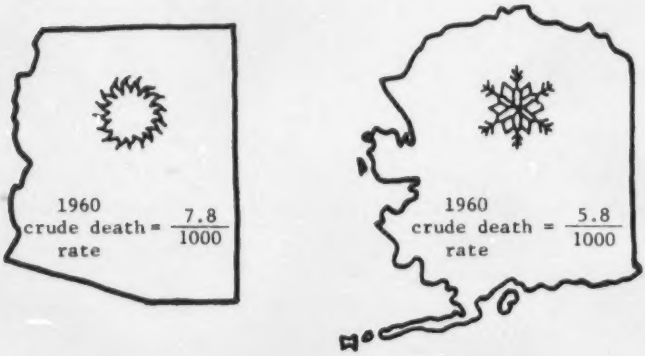
This package is intended to teach you (1) the *conditions* necessary for recognizing such potentially misleading situations; and (2) how to cope statistically with such situations through a procedure which will remove the effects of additional factors (such as age) on the comparison of interest. In general, we call this procedure *rate adjustment*. Rate adjustment uses several different methods; in this package, you will learn how to use the *direct method*.

This package is divided into four sections as follows:

1. When to adjust rates.
2. What adjusted rates do and how to compute the directly adjusted rate.
3. How to interpret adjusted rates.
4. How to compute and interpret directly adjusted rates when you are adjusting for factors other than age.

1. WHEN TO ADJUST RATES

Briefly examine the illustration which compares the overall mortality of two states in 1960.



You might have guessed that the two states pictured above are Arizona (on the left) and Alaska (on the right). If you were a health researcher interested in the effect of climatic conditions on mortality, you might decide to study deaths in these two states. This would allow the comparison of mortality in a cold, damp climate with mortality in a hot, dry one.

Look at the crude death rates for 1960 in these two states:

$$\text{Alaska's crude death rate} = \frac{1313\ (\text{deaths})}{226167\ (\text{population})} = .0058 \text{ or } 5.8 \text{ per } 1000.$$

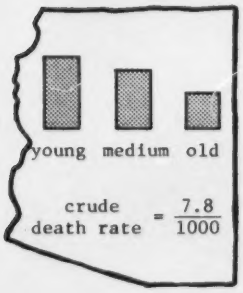
$$\text{Arizona's crude death rate} = \frac{10121\ (\text{deaths})}{1302161\ (\text{population})} = .0078 \text{ or } 7.8 \text{ per } 1000.$$

You might be surprised, considering the climate, to find that Alaska has the smaller rate.

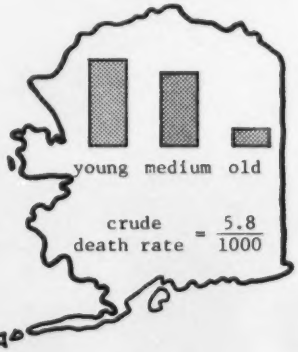
Question 1. What state would you have expected to have had the *higher* rate if climatic conditions were generally associated with mortality?

A little knowledge of the populations of these states might cause you to adjust your interpretation. Look at the age structures of the two states as represented in the following diagram.

ARIZONA



ALASKA



Alaska, a newer state, has tended to attract a younger population. The dry, warm climate of Arizona has, on the other hand, attracted many older persons. In fact, the difference in crude rates can perhaps be at least partially explained by the simple fact that Arizona has an older population than Alaska. We should consequently expect relatively more deaths in Arizona simply because there are relatively more old people there, and old people are at a high risk of dying.

3

The presence of a variable such as *age* in this situation is one of the conditions necessary for computation of adjusted rates. We call such a variable a *confounding factor* because it confounds or blurs the comparison of interest. In other words, the difference we have observed in crude rates can be explained at least partly by the difference in age structures.

Question 2. True or False. If the two crude rates had been exactly the same, this would give strong evidence to suggest that climatic factors have no effects on overall mortality.

A. Four Conditions

Although the presence of a (1) *confounding factor* is the primary condition for rate adjustment, there are in all, *four basic conditions* for rate or proportion adjustment:

(2) You are interested in a *comparison* (not a single population).

(3) The event or characteristic of interest (in this case death) is defined for purpose of analysis as a *rate* (e.g., death rate) or *proportion* (not the mean of a continuous variable like blood pressure where other kinds of adjustment are sometimes used).

(4) Your comparison involves *overall* rates (not specific rates).

Note that all four conditions have to be satisfied to justify adjusting rates. And this is certainly true for the example we have been considering.

Question 3. True or False. If you had been interested in comparing mean number of physicians per county between the two states (instead of comparing death rates), you would need to do rate adjustment.

Question 4. True or False. If for the above example you were *only* interested in comparing the mortality rates for persons in the age group 55-64, you would need to adjust rates with respect to age. Give two reasons for your answer.

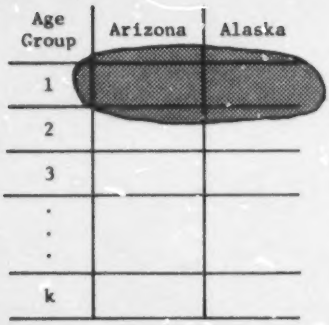
B. Should Overall Rates Ever Be Used at All?

An epidemiologist named Woolsey (1959), expressing an opinion shared by several other researchers, has pointed out that "specific rates are essential because it is only through the analysis of specific rates that an accurate and detailed study can be made of the variation among population classes."

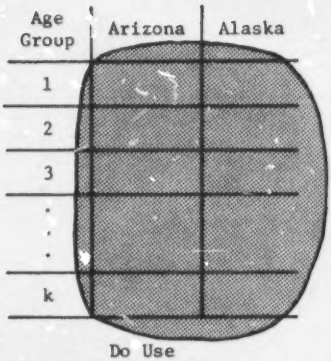
Nevertheless an overall rate can be quite useful as a *convenient summary* of the information in an entire schedule of specific rates. This is essentially because making interpetations can become difficult when the number of specific rates is large. Also, a single rate is especially convenient when additional variables of interest need to be brought into analysis at a later stage.

However, there certainly are situations when use of an overall rate would be inappropriate or at least of questionable value. The criteria for judging when an overall rate should not be used are:

> Interest in a Specific Group. The comparison of interest is clearly restricted to a specific group. (Note, again, that if your specific group is still very broadly defined, then *age* may still be a confounding factor and you may still need to adjust your rates.)

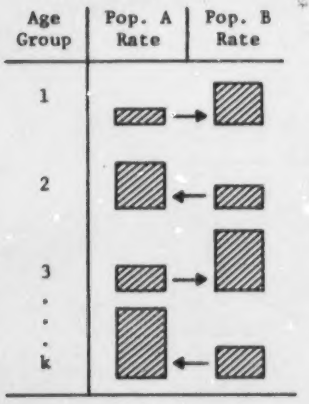


Don't Use Overall Rate

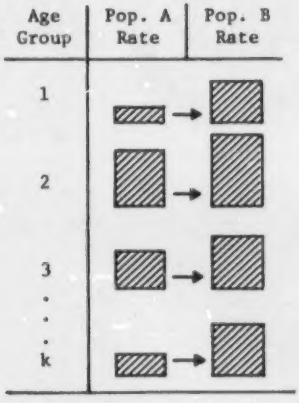


Do Use Overall Rate

Inconsistency. There is noticeable inconsistency in the direction of age-specific differences, i.e., specific rates are noticeably higher for one population at certain ages but noticeably lower for this population at other ages. (Note that when this reason is valid, no single overall rate for each population would pick up the age-specific differences. Rather, use of an overall rate would tend to mask such differences.)



Don't Use Overall Rate



Do Use Overall Rate

The following example illustrates use of the second criterion. Look at the following table of age-specific death rates for Alaska and Arizona in 1960:

ARIZONA (1960)	
Age	Death Rate/1000
< 1	33.9
1-4	1.8
5-14	0.5
15-24	1.5
25-34	1.9
35-44	> 3
45-54	7.7
55-64	17.5
65-74	35.9
75-84	78.3
85+	165.0
	7.8

ALASKA (1960)	
Age	Death Rate/1000
<1	43.1
1-4	2.1
5-14	0.9
15-24	1.4
25-34	1.8
35-44	3.9
45-54	9.1
55-64	16.4
65-74	39.8
75-84	105.6
85+	142.9
	5.8

Note that there is only one age-specific category (i.e., 85+) in the above data for which the rate for Arizona is noticeably higher than the corresponding rate for Alaska. Consequently, it is reasonable to conclude that the direction of the age-specific differences is more or less consistent for these data so that using an overall rate would be meaningful.

Question 5. Suppose you wished to compare two populations with the following mortality rates:

	Population A	Population B
Young	13.2/1000	10.3/1000
Old	9.5/1000	15.9/1000

True or False. An overall rate is appropriate here.

7

C. Review Questions

Suppose you were interested in comparing typhoid fever attack rates resulting from an epidemic in two adjacent communities. Suppose also that you knew that race was associated with typhoid attack rate in the general population. Further suppose the data look like this:

	Community 1 Attack Rate/1000	Community 2 Attack Rate/1000
Blacks	2.18	1.96
Whites	8.99	8.81
Overall Crude Rates	5.04	7.80

1. Can you list the four conditions for rate adjustment (or give a general description that covers these 4 conditions)?
   a. _____ c. _____
   b. _____ d. _____

2. The confounding factor in the above example is _____?

3. Circle as many of the following terms as are appropriate that describe the characteristic of primary interest (i.e., the variable about which a comparison is desired)?
   a. confounding factor c. attack rate
   b. death rate d. mean of a continuous variable

4. a. Are race-specific rates consistently higher for one community over the other community? Yes ____ No ____
   b. Does your answer to 4a support the use of an overall rate? Yes ____ No ____

5. How many of the four conditions for rate adjustment are satisfied?

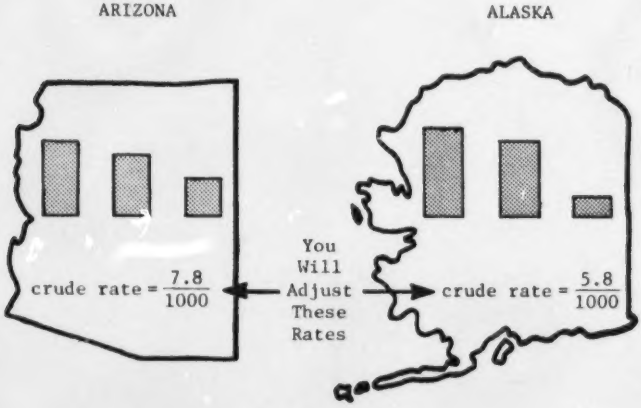
6. True or False. If the rate for Blacks in Community 2 were 5.2 instead of 1.96, you should adjust rates. Why?

Check your answers on page 28.

2. WHAT ADJUSTED RATES DO AND HOW TO COMPUTE THE DIRECT RATE

A. What Adjusted Rates Do

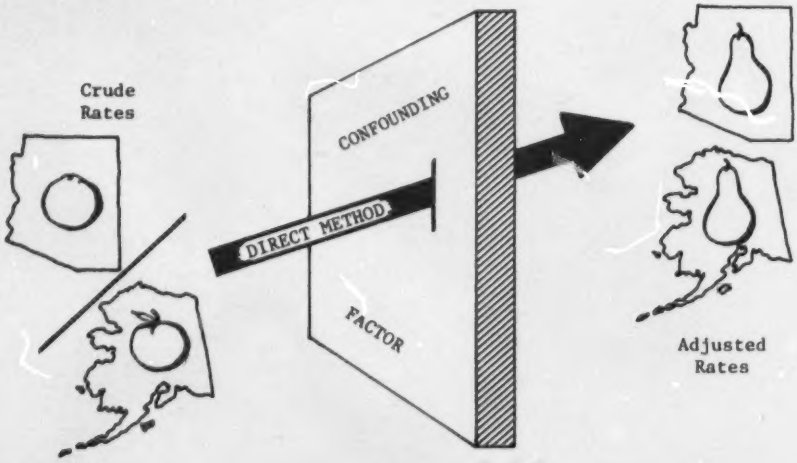
Now that you have learned *when* to adjust rates, you are ready to learn how to perform this adjustment. Let us return to the Arizona-Alaska example:



Recall that this example involves a comparison of *overall rates* in order to draw conclusions about the possible effects of climate on the force of mortality. In making this comparison, we must *remove* the effect of the confounding factor (age), so that any difference in rates that is actually found can not be explained by age differences in the two states.

The method of adjustment to be treated in this program is called the *direct method*. The *direct* method gets its name because it faces the problem of adjustment head on. (Other methods of adjustment are discussed in the references.) It does this by forcing the comparison of the two populations to be made on the basis of a *common* age distribution. In other words, the confounding

factor is directly removed by the *substitution* of a common age distribution for the separate age distributions so that both populations may be compared as if they had the same age structure.



From the above illustration, it can be seen that because of the confounding factor, comparing crude rates is like comparing different kinds of fruits (e.g., orange and apple), whereas comparing adjusted rates is like comparing two fruits of the same kind (e.g., two pears). The direct method removes the confounding factor by substituting a common age distribution for the separate age distributions of the two populations.

The *two basic pieces of information* required for this task are:

a. The (age) *specific* (death) *rates* (or *proportions*) for each test population.
b. A *standard* population.

The information for this example is presented in the two tables on the following page. The age-specific rates for Arizona and Alaska are given in Table 1. A standard population (1960 U.S. standard million) is given in Table 2. 10

TABLE 1

Populations at Risk (PAR) and Age-Specific Death Rates (p) for Arizona and Alaska by Age (1960)

AGE	ARIZONA		ALASKA	
	PAR	p	PAR	p
<1	34599	.0339	7101	.0431
1-4	132367	.0018	27092	.0021
5-14	285830	.0005	46110	.0009
15-24	186789	.0015	40722	.0014
25-34	169878	.0019	39672	.0018
35-44	173029	.0033	31981	.0039
45-54	136573	.0077	18957	.0091
55-64	92871	.0175	9146	.0164
65-74	63634	.0359	3745	.0398
75-84	22499	.0783	1354	.1056
85+	4092	.1650	287	.1429
TOTAL	1302161	.0078	226167	.0058

TABLE 2

Standard Million Population of United States by Age (1960)

AGE	PAR
<1	22883
1-4	89812
5-14	196727
15-24	133591
25-34	126559
35-44	133515
45-54	114381
55-64	92650
65-74	60158
75-84	24933
85+	4791
TOTAL	1000000

Question 6. Look at Tables 1 and 2 to answer the following:

a. What is the 65-74 death rate for Arizona?

b. What is the 35-44 population-at-risk for the standard?

c. What is the 35-44 death rate for the standard?

Notice that the difference between the kind of information given in Table 1 and Table 2 is that Table 2 does not contain any age-specific or even total crude death rates. This was a purposeful omission because *standard rates* are *not needed* (and sometimes not even known) for the computation of the direct rate. Actually, the age-specific PAR's given in Table 1 are not formally needed either and we have put them in only because this is usually done for completeness and to give a sense of the difference in age structures.

The *standard population* is that *common* distribution referred to above whose primary purpose is to serve as a *reference group* or *stand-in* (*substitute*) for the different age distributions of Arizona and Alaska. The choice of the standard depends upon the particular situation and is in some sense always arbitrary. The standard used here was the U.S. 1960 standard million because this latter group was a reasonable common denominator for the populations of the two states. This choice also has the advantage of being equally good for any other state we later decide to compare with both Arizona and Alaska. Generally, the standard is chosen to agree as closely as possible with the populations of interest. (Often, indeed, the average of the two populations is used as the standard.) Nevertheless, as you would hope, the choice of the standard usually (although not always) does not affect the direction of the results of your comparison.

Question 7. What choice of standard population would you suggest for comparing rates in England and Wales in 1970?

Question 8. Which of the following populations would be *most* appropriate for comparing 1974 death rates in two North Carolina counties?

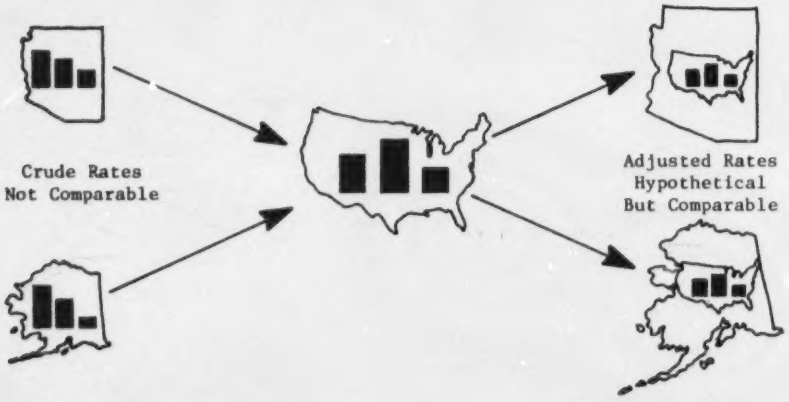
a. 1974 U.S. pop. b. 1974 N.C. pop. c. 1960 N.C. pop.

Question 9. Which of the following populations would be *least* appropriate as a standard for comparing 1974 death rates of whites and blacks in a given N.C. county?

a. 1974 N.C. pop.
b. 1974 pop. of the given county
c. 1970 pop. of the given county
d. 1970 U.S. pop.

To summarize, the basic idea in computing a direct rate for a test population (e.g., Alaska) is to compute what the (hypothetical) crude rate would be for the test population if it had the same age structure as the standard (e.g., United States). When this is done for two test populations using the same standard, the confounding factor is removed because the two populations are thus being treated as if they had the same age structure.

Note that since neither Alaska nor Arizona actually has the same age structure as the United States, their direct rates (using the same standard) are hypothetical. However, although the adjusted rates are hypothetical, they are nevertheless comparable.



B. How To Compute the Direct Rate

You will now learn how to compute the direct adjusted rates for Arizona and Alaska using Tables 1 and 2. The *procedure* for computing the direct rate for any given test population involves three steps:

Step 1: Compute expected cases for each specific group.

Step 2: Compute total expected cases.

Step 3: Compute direct rate.

Now, in the Arizona-Alaska example, the *cases* we are considering are *deaths*. The total *expected* cases refer to the hypothetical number of deaths in the test population (e.g., Alaska or Arizona) that would be expected if the test population had the same age structure as the standard. To get the total expected cases, you must sum over all groups the expected cases for each specific group.

Tables 1 and 2 are repeated here because they are both needed to compute Step 1:

TABLE 1

Populations at Risk (PAR) and Age-Specific Death Rates (p) for Arizona and Alaska by Age (1960)

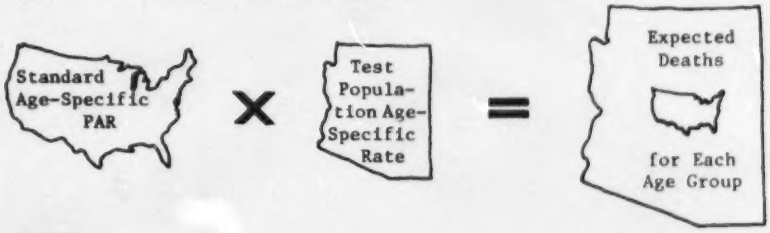
AGE	ARIZONA		ALASKA	
	PAR	p	PAR	p
<1	34599	.0339	7101	.0431
1-4	132367	.0018	27092	.0021
5-14	285830	.0005	46110	.0009
15-24	186789	.0015	40722	.0014
25-34	169878	.0019	39672	.0018
35-44	173029	.0033	31981	.0039
45-54	136573	.0077	18957	.0091
55-64	92871	.0175	9146	.0164
65-74	63634	.0359	3745	.0398
75-84	22499	.0783	1354	.1056
85+	4092	.1650	287	.1429
TOTAL	1302161	.0078	226167	.0058

TABLE 2

Standard Million Population of United States by Age (1960)

AGE	PAR
<1	22883
1-4	89812
5-14	196727
15-24	133591
25-34	126559
35-44	133515
45-54	114381
55-64	92650
65-74	60158
75-84	24933
85+	4791
TOTAL	1000000

Step 1: Find the expected number of deaths for each age group by *multiplying* the standard population PAR by the test population rate for each age-specific group, e.g.,



For an example of Step 1, looking back at Tables 1 and 2 to get the data, you can multiply as follows to get the expected deaths for ages 5-14 in Arizona:

196727 × .0005 = 98 (rounded off).

(Note that conventions about rounding off vary. For simplicity we shall use rounding to the nearest whole number in the Arizona-Alaska example, but in later examples we will carry one or two decimal places—even though that will seem to involve thinking about fractions of deaths or of disease cases. In practice it is usually sensible to carry at least one more decimal place in intermediate calculations than you plan to use in the final result.)

Question 10. Using Tables 1 and 2, find the expected deaths for ages 55-64 in Arizona.

_____ × _____ = _____.

Question 11. Looking back at Tables 1 and 2, how many expected death calculations must be carried out in order to compute direct rates for Arizona and Alaska?

In order to perform more conveniently all of the necessary computations required for the direct method, the basic

information of Tables 1 and 2 is rearranged into the following table. The unnecessary information is left out of this composite table and most, but not all, of the calculations have been performed.

TABLE 3

Standard Computing Format for Direct Rate Adjustment

AGE	STANDARD PAR	ARIZONA		ALASKA	
		p	E	p	E
	Col. 1	Col. 2	Col. 3 = 1×2	Col. 4	Col. 5 = 1×4
< 1	22883	.0339		.0431	
1-4	89812	.0018		.0021	189
5-14	196727	.0005	98	.0009	177
15-24	133591	.0015	200	.0014	187
25-34	126559	.0019	240	.0018	228
35-44	133515	.0033	441	.0039	521
45-54	114381	.0077	881	.0091	1041
55-64	92650	.0175	1621	.0164	
65-74	60158	.0359	2160	.0398	2394
75-84	24933	.0783	1952	.1056	2633
85+	4791	.1650	791	.1429	685
Total	1000000				

The main difference between Table 3 and Tables 1 and 2 is that space is provided in Table 3 for the results of Step 1, expected deaths (E). You will be asked to construct a table like this later.

Question 12. Fill in the remaining blanks of expected deaths for Arizona in Table 3. You may use the following space for calcuations:

Arizona <1: _____ × _____ = _____.

Arizona 1-4: _____ × _____ = _____.

Step 2: Compute total expected deaths by *adding* expected deaths over all age-specific groups.

Step 2 is easy because all you have to do is add the expected deaths computed in Step 1 over all age groups (separately for Arizona and Alaska).

Question 13. Compute the total expected deaths for Arizona.

Step 3: Compute the direct rate by *dividing* the total expected deaths by the total standard population.

Step 3 yields the two (direct) adjusted rates which are comparable, though hypothetical. Thus, the direct adjusted rate for Arizona:

9322/1000000 = .009322

or 9.3 per 1000 if we round to one decimal place.

Question 14. Now perform Steps 1, 2 and 3 as needed to complete Table 3 for Alaska and arrive at the direct rate for Alaska.

C. Review Questions

Suppose you are given the following data for comparing the coronary heart disease (CHD) rates of white male *nonfarm* workers 40 and over with those of while male *farm* workers 40 and over in a certain county in Georgia.

TABLE 4

Age	Nonfarmers		Farmers		Standard (combined groups)	
	PAR	p(CHD)	PAR	p(CHD)	PAR	p
40-44	72	.125	37	.000	109	.083
45-49	158	.089	91	.055	249	.076
50-54	79	.177	76	.079	155	.129
55-59	47	.277	43	.186	90	.233
60+	4	.500	2	.000	6	.333
Total	360	.145	249	.076	609	.117

1. Rearrange Table 4 in standard format so that adjusted rates can be conveniently computed and nonessential information is left out. (You need not fill in the cells that require multiplication, summation or division.) Try to do this without looking back to page 17.

Age					

2. Use the above rearranged table to compute adjusted rates for nonfarmers and farmers.

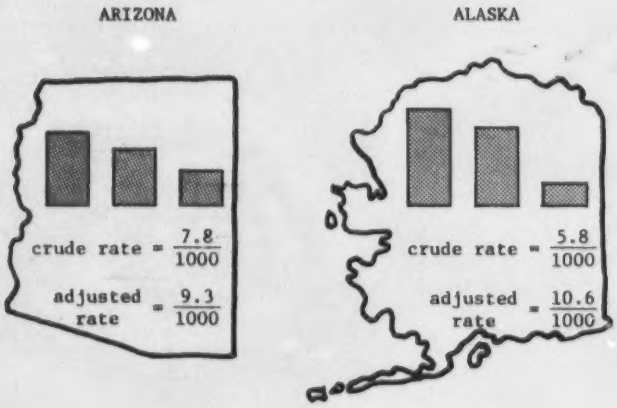
Nonfarmers: _____ _____ or _____ per 100.

Farmers: _____ or _____ per 100.

Check your answers on page 29.

3. INTERPRETING ADJUSTED RATES

Again let us return to our Arizona-Alaska problem to interpret what has been achieved by adjustment.



Looking at the two adjusted rates in the above picture, you notice that a very interesting change has occurred. The adjusted rate for Alaska (10.6/1000) is higher than the adjusted rate for Arizona (9.3/1000)! This is a *reversal* from the earlier crude rates (5.8/1000 for Alaska and 7.8/1000 for Arizona). Thus, when the differences in age structure of the populations in Alaska and Arizona are removed, the force of mortality in Alaska is actually higher than that in Arizona.

Question 15. Using the adjusted rates for Alaska and Arizona, which state appears to be better off with regard to mortality?

Question 16. Which of the following do you think best describes the likelihood of your getting a similar reversal if a standard other than U.S. 1960 was used?

a. impossible to get reversal with another standard
b. certain to get reversal with another standard
c. possible, but generally unlikely to get reversal with another standard
d. generally likely to get a reversal with another standard.

You may wish to review what you have learned by working through another example. Furthermore, you may be interested in an example for which the confounding factor is not age.

4. REVIEW USING A CONFOUNDING FACTOR OTHER THAN AGE

Let us return to the data on typhoid fever attack rates resulting from an epidemic in two adjacent communities:

TABLE 5

Race	Community 1		Community 2		Standard (Combined Communities)
	PAR	p	PAR	p	PAR
Black	2757	.00218	1020	.00196	3777
White	2002	.00899	5901	.00881	7903
Total	4759	.00504	6921	.00780	11680

You should have previously (page 8) concluded that *race-adjustment* was appropriate for comparing the attack rates of the two communities controlling for the confounding factor race.

Question 17. Rearrange Table 5 in standard form leaving out extraneous information and compute the direct race-adjusted rates. Use the space below:

Race					
____	____	____	____	____	____
____	____	____	____	____	____
Total:	____	____	____	____	____

Adjusted rate for Community 1:
Adjusted rate for Community 2:

Question 18. Using the adjusted rates just computed, which community has the higher attack rate?

Question 19. True or False. The community with the higher crude rate also had the higher direct rate.

Question 20. Did the adjustment process widen or narrow the difference in rates between the two communities?

Every example in this module has adjusted rates for a single confounding factor. These techniques, however, can be extended to adjust for several confounding factors, for example, age and race simultaneously. All that is needed is rates specific to each subgroup (e.g., death rates for white males age 15-24) and a standard population classified into such subgroups.

To test your knowledge take the post test that follows. If you get less than 90 percent on this test, you should either review this module or read from another source (see Section 11, References).

5. POST TEST

(At least 90 percent is the expected performance level.)

I. The following table presents mortality data for year X on two hypothetical communities (A and B) in North Carolina, giving the population (PAR) and the death rates (p) for each community and for a combined standard (A + B) within each of three age groups:

TABLE Ia

Age	Community A		Community B		Standard (A+B)	
	PAR	p	PAR	p	PAR	p
Young	2000	.0020	1000	.0010	3000	.0017
Medium	2000	.0050	2000	.0050	4000	.0050
Old	2000	.0100	3000	.0090	5000	.0094
Total	6000	.0057	6000	.0063	12000	.0060

Points

1. a. What conditions need to be satisfied to justify the use of rate adjustment for this example?

   i. _____

   ii. _____

   iii. _____

   iv. _____ 4 pts. ___

   b. Which of these conditions are satisfied for this example? _____ 2 pts. ___

   c. Assuming that you are not interested in only one specific age group, what should you check to determine whether use of overall rates are appropriate?

   _____

   _____

   _____ 3 pts. ___

23

Points

d. Give the result of this check for the above example.

_____

_____ 3 pts. ___

2. a. Which of the following standard populations would be least appropriate for use in adjustment of the above data?

   i. Pooled communities (A + B) for Year X
   ii. U.S. population for Year X + 2
   iii. Pooled communities for Year X + 2
   iv. N.C. population for Year X 3 pts. ___

   b. Which of the following characteristics are appropriate for describing the standard population in this example?

   i. Stands-in for the age distribution of both A and B
   ii. Should be chosen to resemble A and B as much as possible
   iii. Standard rates are not needed for computation of the direct rate 3 pts. ___

   c. Which of the following characteristics correctly describe the adjusted rate for Community A?

   i. The crude rate for Community A if this community had the age-specific rates of the standard
   ii. A hypothetical rate
   iii. Comparable to the adjusted rate of B if the same standard is used 5 pts. ___

3. a. Complete the following table for calculation of the direct rate:

TABLE Ib

Age	Standard (A + B)	Community A		Community B	
		p	Expected Deaths	p	Expected Deaths
Young	3000	.0020	_____	.0010	_____
Medium	4000	.0050	_____	.0050	_____
Old	5000	_____	50	.0090	_____
Total	12000	_____	_____	_____	_____

Points

8 pts. ___

b. Using your results in (3a) compute the direct adjusted rates for each community.

Community A: _____

Community B: _____ 8 pts. ___

c. Based on your adjusted rates, which community is worse off? _____ 3 pts. ___

d. True or False. You would have made a different conclusion if you had only used the crude rates for Communities A and B. 3 pts. ___

II. Table IIa gives incidence rates of mongolism (per *live birth*) specific to birth order for two maternal age groups in Michigan for the period 1950-1964.

TABLE IIa

Birth Order	Maternal Age 20-24		Maternal Age < 20		Michigan	
	PAR	p	PAR	p	PAR	p
1	330000	.00043	230000	.00047	731000	.00056
2	327000	.00046	72000	.00035	725000	.00068
3	176000	.00040	15000	.00020	569000	.00083
4	69000	.00038	2000	.00044	358000	.00115
5+	31000	.00026	500	.00000	443000	.00167
Total	933000	.00043	319500	.00043	2826000	.00090

25

Points

4. a. What is the confounding factor in this example?

_____ 3 pts. ___

b. How many birth order groups contain higher specific rates for maternal age group < 20 than for the 20-24 group? _____ 3 pts. ___

c. True or False. Your answer in (4b) gives support for rate adjustment. _____ 3 pts. ___

5. a. What two basic pieces of information have been provided which are necessary for computation of adjusted rates?

i. _____

ii. _____ 5 pts. ___

b. Whether or not you think it is appropriate in this case to use an overall rate, rearrange the above data into standard computing format and fill in all the blanks in the table below.

TABLE IIb

Birth Order	Michigan Standard PAR	Maternal Age 20-24		Maternal Age < 20	
1	731000				
2	725000				
3	569000				
4	358000				
5+	443000				
Total	2826000				

18 pts. ___

	Points
c. Compute the adjusted rates for both maternal age groups. 20-24: _____ < 20: _____	8 pts. ___
6. a. Which group is of higher risk for mongolism births? _____	5 pts. ___
b. True or False. There was a reversal in the comparison when going from crude to adjusted rates.	5 pts. ___
c. State the one reservation that you should have regarding the conclusions you have reached in (6a) and (6b). _____ _____	5 pts. ___
	TOTAL SCORE: ☐

6. RECOMMENDED FOLLOW-UP ACTIVITIES

It is recommended that you now study the indirect method of rate adjustment and its comparison to the direct method. Then you may want to continue with statistical inference for comparing adjusted rates.

7. APPENDIX: A STANDARD MILLION

A standard million is a population (PAR) whose numbers in specific categories have been changed in order to total 1 million, while remaining the same relative proportions. The process of revision to total 1 million amounts to finding the proportionate part of the total population in each specific group (analagous to the

percentage in each specific group) and multiplying the proportion by 1,000,000. An example demonstrating this process for the 1960 U.S. population categorized by age is given in Table A below.

TABLE A

Computation of the 1960 U.S. Standard Million

Age	1960 U.S. PAR	proportionate part = $\frac{\text{age-specific PAR}}{180,325,775}$	Standard Million PAR = 1,000,000 × proportionate part
< 1	4,126,403	.022883	22,883
1-4	16,195,413	.089812	89,812
5-14	35,474,882	.196727	196,727
15-24	24,089,957	.133591	133,591
25-34	22,821,888	.126559	126,559
35-44	24,076,192	.133515	133,515
45-54	20,625,775	.114381	114,381
55-64	16,707,225	.092650	92,650
65-74	10,848,086	.060158	60,158
75-84	4,496,032	.024933	24,933
85+	863,922	.004791	4,791
Total	180,325,775	1.000000	1,000,000

8. ANSWERS TO REVIEW SECTION QUESTIONS

1C. Review Questions

1. a. rate or proportion
   b. comparison
   c. overall rate
   d. confounding factor

2. The best answer from the information you are given is *race*. Although it is possible that age might also be a factor, the example does not consider age.

3. Only one of the four descriptions is correct, namely c.

4. Yes to both questions.

5. All four.

6. False, because the race-specific rates would not vary in the same direction for each race. Thus an overall rate would not be meaningful and this condition for adjustment would not be met.

If you got all the answers on the review right you are ready to proceed to Section 2. If you missed more than one you ought to reread Section 1 before proceeding.

2C. Review Questions

1. Your rearranged table should look as follows:

Age	Standard PAR	Nonfarmers		Farmers	
		p	exp. cases	p	exp. cases
40-44	109	.125	____	.000	____
45-49	249	.089	____	.055	____
50-54	155	.177	____	.079	____
55-59	90	.277	____	.186	____
60+	6	.500	____	.000	____
Total	609		____		____

2.

Age	Standard PAR	Nonfarmers		Farmers	
		p	exp. cases	p	exp. cases
40-44	109	.125	13.625	.000	0
45-49	249	.089	22.161	.055	13.695
50-54	155	.177	27.435	.079	12.245
55-59	90	.277	24.930	.186	16.740
60+	6	.500	3.000	.000	0
Total	609		91.151		42.68

Adjusted Rate for Nonfarmers: $\frac{91.151}{609} = 0.1497$ or 14.97 per 100.

Adjusted Rate for Farmers: $\frac{42.68}{609} = 0.0701$ or 7.01 per 100.

(Note that we have earlier expressed rates as number per thousand but we are expressing these rates as number per hundred. The choice of such a base is really arbitrary. The guiding principle is usually that the smallest rate has a single digit to the left of the decimal point.)

9. ANSWERS TO IN-TEXT QUESTIONS 1-20

1. Alaska, because of its colder, damper climate, which you would expect to make it have a higher mortality rate.

2. False. Though there would be some evidence to suggest that there are no climatic effects overall, the comparison of interest is confounded by the factor age. If no difference in crude rates is observed, this may be entirely due to difference in age structures between Alaska and Arizona. Similarly, any large difference (such as the one we observed in the illustration) might also be explained entirely by the age factor. Also, there may be other confounding factors such as the number or quality of medical care facilities that could explain any observed difference or mask any true difference.

3. False, because your event of interest is not defined as a rate but rather as a mean.

4. False, because (a) you would not be interested in comparing *overall* rates, and (b) because age would not be a confounding factor. Note, however, that the age group 55-64 may not be restricted narrowly enough so that within this age group, there may still be confounding due to age.

5. The answer here should be False since the comparison of rates within age-specific groups (broadly classified into old and young) differs greatly and in a different direction depending on the age group.

6. a. .0359 b. 133515 c. not given.

7. Either total Great Britain 1970 or England and Wales 1970 would be good. U.S. 1960 would not be as good as it would be further removed in time and place.

8. (b.) since 1974 N.C. population is more closely related to the two counties than the other two.

9. Since 1974 U.S. population is the least related to the population of the two groups of interest, the answer is (d.).

10. $92650 \times .0175 = 1621$.

11. 22, since there are 11 age groups in Arizona and 11 in Alaska.

12. Arizona < 1: $22883 \times .0339 = 776$
Arizona 1-4: $89812 \times .0018 = 162$.

13. 9322, which is obtained by summing all expected deaths in column 3 of Table 3.

14. .010560 or 10.6 per 1000 if we round to one decimal place.

15. Your answer should be Arizona, since it has the lower adjusted rate.

16. The correct answer is (d.), because usually (though not always) the results will be comparatively the same regardless of the standard chosen, especially if the standard reasonably relates to the populations being compared.

17. Your answer should be as in the following table:

Race	Standard PAR	Community 1		Community 2	
		p	exp. deaths	p	exp. deaths
Black	3777	.00218	8.234	.00196	7.403
White	7903	.00899	71.048	.00881	69.625
Total	11680		79.282		77.028

$$\text{Adjusted rate} = \frac{\text{total expected deaths}}{\text{total standard population}}$$

Community 1: $\frac{79.282}{11680} = .00679$ or 6.79/1000

Community 2: $\frac{77.028}{11680} = .00659$ or 6.59/1000.

18. Your answer is Community 1.

19. False.

20. Your answer should be narrow, since the difference between crude rates is 7.80/1000 - 5.04/1000 = 2.76/1000 whereas the difference between adjusted rates is 6.79/1000 - 6.59/1000 = 0.20/1000. Thus the adjustment process narrowed the difference considerably.

10. ANSWERS TO POST TEST

I. 1. a. (i) rate or proporation (iii) overall rate
(ii) comparison (iv) confounding factor

b. all of them

c. Check to see whether differences in specific rates are consistently in the same direction over all specific groups.

d. Community A's specific rates are as high or higher than those of Community B in every case, if rates are consistent. Thus an overall rate is appropriate.

2. a. ii

b. All of them

c. All but (i)

3. a.

TABLE Ib

Age	Standard (A+B)	Community A		Community B	
		p	Expected Deaths	p	Expected Deaths
Young	3000	.0020	6	.0010	3
Medium	4000	.0050	20	.0050	20
Old	5000	.0100	50	.0090	45
Total	12000		76		68

b. Community A: $\frac{76}{12000} = .00633$ or 6.33/1000

Community B: $\frac{68}{12000} = .00567$ or 5.67/1000.

c. Community A

d. True

II. 4. a. Birth order

b. Two

c. False

5. a. (i) (birth order) specific rates
(ii) standard population

b. TABLE IIb

Birth Order	Michigan Standard PAR	Maternal Age Group 20-24		Maternal Age Group < 20	
		p	Exp. Cases	p	Exp. Cases
1	731000	.00043	314.33	.00047	343.57
2	725000	.00046	333.50	.00035	253.75
3	569000	.00040	227.60	.00020	113.80
4	358000	.00038	136.04	.00044	157.52
5+	443000	.00026	115.18	.00000	000.00
Total	2826000		1126.65		868.64

c. Age 20-24: $\text{Adjusted Rate} = \frac{1126.65}{2826000} = .000399$ or 39.9/100,000

Age < 20: $\text{Adjusted Rate} = \frac{868.64}{2826000} = .000307$ or 30.7/100,000.

6. a. 20-24 maternal age group.

b. True (crude rates were equal, adjusted rates different).

c. Since one maternal age group does not have consistently higher birth-order specific rates than the other group, use of overall rates is questionable, because overall rates mask birth order specific differences.

11. REFERENCES

Colton, T. 1974. Statistics in Medicine. Chapter 2. Boston: Little Brown & Company.

Fleiss, J.L. 1973. Statistical Methods for Rates and Proportions. Chapter 13. New York: John Wiley & Sons, Inc.

Remington, Richard D. and Schork, M. Anthony. 1970. Statistics with Application to the Biological and Health Sciences. Chapter 13. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Woolsey, T.D. 1959. Adjusted death rates and other indices of mortality. Chapter 4 in F.E. Linder and R.D. Grave, Vital Statistics in the United States 1900-1940. Washington, D.C.: U.S. Government Printing Office.

STUDENT FORM 1

Request for Help

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

Student: If you have trouble with a specific part of this unit, please fill out this form and take it to your instructor for assistance. The information you give will help the author to revise the unit.

Your Name_____ Unit No._____

Page_____
○ Upper
○ Middle
○ Lower

OR

Section_____
Paragraph_____

OR

Model Exam
Problem No._____
Text
Problem No._____

Description of Difficulty: (Please be specific)

Instructor: Please indicate your resolution of the difficulty in this box.

○ Corrected errors in materials. List corrections here:

○ Gave student better explanation, example, or procedure than in unit. Give brief outline of your addition here:

○ Assisted student in acquiring general learning and problem-solving skills (not using examples from this unit.)

Instructor's Signature_____

Please use reverse if necessary.

STUDENT FORM 2

Unit Questionnaire

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

Name_____Unit No._____Date_____

Institution_____Course No._____

Check the choice for each question that comes closest to your personal opinion.

1. How useful was the amount of detail in the unit?

___Not enough detail to understand the unit
___Unit would have been clearer with more detail
___Appropriate amount of detail
___Unit was occasionally too detailed, but this was not distracting
___Too much detail; I was often distracted

2. How helpful were the problem answers?

___Sample solutions were too brief; I could not do the intermediate steps
___Sufficient information was given to solve the problems
___Sample solutions were too detailed; I didn't need them

3. Except for fulfilling the prerequisites, how much did you use other sources (for example, instructor, friends, or other books) in order to understand the unit?

___A Lot ___Somewhat ___A Little ___Not at all

4. How long was this unit in comparison to the amount of time you generally spend on a lesson (lecture and homework assignment) in a typical math or science course?

Much	Somewhat	About	Somewhat	Much
___Longer	___Longer	___the Same	___Shorter	___Shorter

5. Were any of the following parts of the unit confusing or distracting? (Check as many as apply.)

___Prerequisites
___Statement of skills and concepts (objectives)
___Paragraph headings
___Examples
___Special Assistance Supplement (if present)
___Other, please explain_____

6. Were any of the following parts of the unit particularly helpful? (Check as many as apply.)

___Prerequisites
___Statement of skills and concepts (objectives)
___Examples
___Problems
___Paragraph headings
___Table of Contents
___Special Assistance Supplement (if present)
___Other, please explain_____

Please describe anything in the unit that you did not particularly like.

Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)